

심혈관계 질환 진단을 위한 정책 기반 워크플로우 시스템

김회영¹⁾, 정우람¹⁾, 윤찬현^{1) 2)}
한국과학기술원 정보통신공학과¹⁾, 전자 전산학부²⁾

Policy based Grid Workflow System for Identification of Heart Disease

Hoeyoung Kim¹⁾, Wooram Jung¹⁾, Chan-Hyun Youn^{1) 2)}
Dept of Information and Communications¹⁾,
School of Engineering and Computer Science²⁾
Korea Advanced Institute of Science and Technology

요 약

본 논문에서는 그리드 기반 워크플로우 시스템을 심혈관계 질환 진단 시스템에 적용하고 그 성능을 평가하였다. 논문은 그리드 기반 워크플로우의 정책을 정의하고 CIM기반의 표준 모델로 구현하였다. 또한 정의된 정책은 기존의 그리드 기반 자원 관리 시스템의 정책과 통합하기 위한 정책 결정자를 제안하였으며 이를 Pegasus와 Kepler기반으로 구현하였다. 응용 및 성능 평가를 위하여 제안하는 시스템은 심혈관계 질환 진단 시스템에 적용 되었으며 사용자의 SLA를 시간과 비용관점에서 효과적인 수행을 만족시킬 수 있음을 실험을 확인하였다.

1. 서론

심혈관계 질환 진단 시스템의 발달에도 불구하고 여전히 심혈관계 질환에 대한 민감도(sensitivity)와 신뢰도에 대하여 한계를 나타내고 있는 것이 현실이다. 이러한 문제를 해결하기 위하여 심혈관계 진단 방법들을 그리드 기술을 이용하여 통합하고 진단 서비스를 의사에게 제공하는 방법이 현재 연구 중에 있다. 이러한 시스템은 각각의 진단 방법을 서로 보완하고, 고성능의 진단을 위하여 제시한 진단 방법들을 통합하여 진단하는데 이를 위한 그리드 워크플로우 기술은 진단의 정확도와 신속성을 높여주는데 사용될 수 있다. [1-3]

그리드 워크플로우 시스템은 사용자가 필요로 하는 다양한 요구사항을 다양하게 고려하여 워크플로우 스케줄링을 하기 위한 연구가 활발히 되어 있지 않다. 대부분의 그리드 워크플로우 시스템은 사용자의 SLA를 고려하지 않고 스케줄링을 하려는 것을 알 수 있다. 따라서 본 논문에서 이에 따라 워크플로우를 위한 사용자 SLA를 정의하고 이를 만족시키기 위한 정책기반 스케줄링 기법을 및 그 시스템 구조를 제안한다. 정책기반 스케줄링은 기 개발된 그리드 자원 관리 시스템

과의 통합을 통해 수행되는데 이때 그리드 워크플로우 정책과 그리드 자원 관리 시스템의 정책과의 정책 조정이 일어나게된다. 현재 본 논문에서 제시된 시스템은 기 개발된 그리드 자원 스케줄러 PQRM(Policy-Quorum based Resource Management)의 기능을 확장하여 페가수스(Pegasus)와 케플러(Kepler)기반의 정책기반 그리드 워크플로우 시스템을 개발하였으며 이를 심혈관계 질환 진단 시스템에 적용하여 시간과 비용관점에서 효율적인 수행을 확인하였다. [4-5]

논문의 구성은 다음과 같다. 2장에서는 그리드 워크플로우 시스템이 적용될 심혈관계 진단 시스템과 기 개발된 그리드 자원 관리 시스템에 대해 설명한다. 3장에서는 그리드 워크플로우 시스템의 구조와 정책 조정자 구조를 나타내며 4장에서는 이를 이용한 실험 결과를 보여준다. 마지막으로 5장에서는 실험결과를 분석하고 의미를 분석한다.

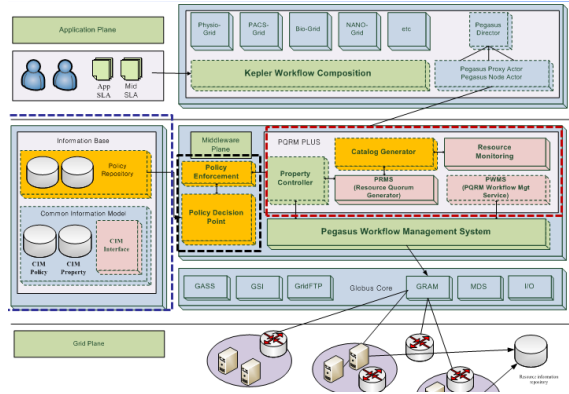
2. 관련 연구

2.1. 협업적 심혈관계 질환 진단 시스템

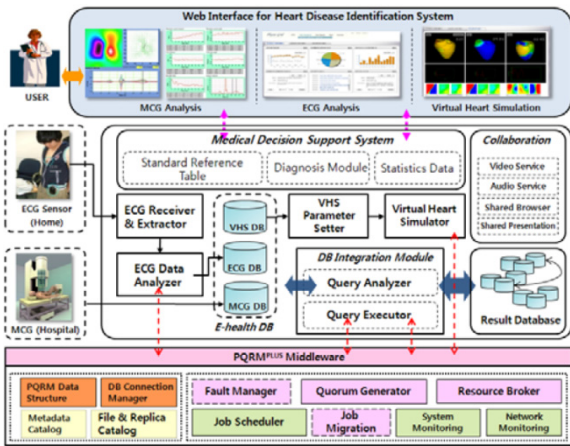
Physio-Grid 시스템은 보다 신속하고 정확한 심혈관계 진단 서비스를 의사와 환자에게 제공하기 위해 기존의 심혈관계 진단 방법(ECG,MCG, Virtual Heart)

들을 그리드 기반으로 통합하여 각각이 가지고 있는 단점 및 한계를 극복하고 유비쿼터스 시대의 e-Health 서비스를 제공해 줄 수 있는 의료 시스템이다. 시스템은 우선 심전도 센서를 통한 데이터를 공중망을 통해 Physio-Grid 시스템 서버 및 DB에 전송되고 처리되며, 또한 로컬 영역에서의 심자도 데이터를 웹을 이용해 모으고, 가상 심장을 통한 심혈관계 질환 진단을 웹 환경 기반으로 사용자에게 제공한다. 최종적인 진단 결과는 해당 사용자에게 웹을 통해 제공함으로써, 고성능의 진단 서비스를 요청한 해당 사용자는 사용자가 원하는 환자의 심혈관계 질환에 대한 진단 서비스를 신속하게 제공받을 수 있다. Physio-Grid시스템은 통합 진단 시스템에서 필수적인 다수개의 데이터 저장소에 분산 배치되어 저장된 심혈관계 질환 관련 각종 진단 데이터에 대한 효율적인 통합 관리 서비스를 실현할 수 있다. 아래의 그림은 PhysioGrid 시스템이 제공하는 서비스를 위한 아키텍처를 나타낸다.

서 시스템은 CIM기반의 Policy저장저장부터 Policy를 꺼내고 이를 PQRM시스템에서 얻은 ARQ에 적용하여 워크플로우를 위한 자원 집합인 AARQ를 얻게 된다. 이렇게 얻어진 자원 집합을 이용하여 워크플로우를 수행하게 된다.



[그림2] 정책 기반 구조도



[그림1] PhysioGrid 시스템 개념도

3. 정책 기반 그리드 워크플로우 시스템

3.1. 시스템 구조

아래의 그림은 정책기반의 워크플로우 시스템의 내부를 나타내는 시스템 구조를 나타낸다. 시스템은 1) Policy를 저장하고 관리하는 CIM기반의 저장소 2) 개발된 그리드 리소스 관리 시스템인 PQRM 3) 그리드 워크플로우의 스케줄링을 맡고 있는 페가수스기반의 시스템 그리고 마지막으로 그리드 워크플로우를 구성하고 실행하기 위한 케플러로 구성되어 있다. 시스템을 이용하고자 하는 사용자는 케플러를 이용하여 심혈관계 질환을 위한 프로세스를 정의하고 이와 관련된 응용 서비스 정책, 워크플로우 정책, 리소스 정책을 정의한다. 이렇게 워크플로우와 정책은 페가수스기반의 시스템으로 스케줄링을 위해 이동되게 된다. 여기

3.2. 그리드 워크플로우 SLA 및 정책

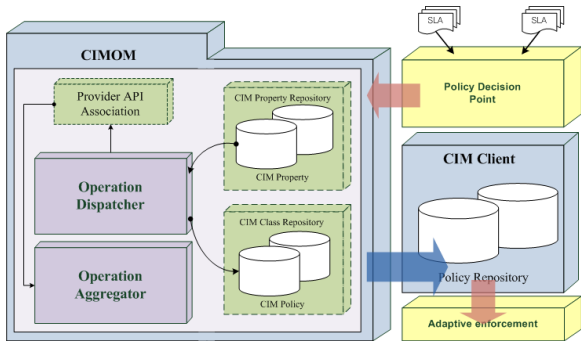
그리드 워크플로우 SLA는 3가지의 정책으로 표현될 수 있다. 이는 응용 정책, 워크플로우 정책, 자원 정책으로 나누어 구성되어 있다. 응용 정책은 현재 총 3가지로 나누어져 있으며 심혈관계 질환 진단을 위한 프로세스의 복잡성 및 정밀도에 따라 3가지로 나누어 생각 할 수 있다. 워크플로우 정책은 워크플로우의 실행 비용의 중요도에 따라 높음, 중간, 낮음의 3단계로 구성하였고 리소스의 정책은 중요도에 따라 6단계로 구성하였다.

	ID	Action
응용 서비스 정책	App-A	ECG 데이터 관리 ECG 분석 서비스 질병 분석 가상 심장 서비스
	App-B	ECG 데이터 관리 ECG 분석 서비스 질병 분석
	App-C	ECG 분석 서비스
워크플로우 정책	WP-A	Enforce ()
	WP-B	Enforce ()
	WP-C	Enforce ()
리소스 정책	RR-A	"Set PQRM Service Level ID : 1001"
	RR-B	"Set PQRM Service Level ID : 1002"
	RR-C	"Set PQRM Service Level ID : 1003"
	RR-D	"Set PQRM Service Level ID : 1004"
	RR-E	"Set PQRM Service Level ID : 1005"
	RR-F	"Set PQRM Service Level ID : 1006"

[표1] 정책 종류

3.3. CIM기반 Policy

CIM은 Policy를 표현하고 저장할 수 있는 표준적인 방법을 제공한다. 이를 통해 Policy의 구성요소인 Condition과 Action을 표현 할 수 있다. CIM은 이와 함께 서로 다른 미들웨어간의 상호 호환성을 위한 기초를 마련해준다. 실제 CIM 기반 Policy는 Kepler Property, Pegasus Property 그리고 PQRM Property로 구분되어 구현된다.



[그림3] CIM기반 Policy

3.4. 정책 결정 방법

```

Algorithm. Policy Decision Scheme
Input: SLA, Workflow, ARQ
Output: CIM based Policy

 $SLA_{app}^k = \{R^k, Pf^k, As^k\}$ 
 $R^k$  : Resource Level Policy
 $Pf^k$  : Workflow Performance Policy
 $As^k$  : Grid Application Policy

START
Action $_{app}^k$  (app properties)
    = Action for  $P_{user}^k$  for  $W_{high}^k$ 
Action $_{app}^k$  (app properties)
    = Action for  $P_{user}^k$  for  $W_{high}^k$ 
Action $_{app}^k$  (app properties)
    = Action for  $P_{user}^k$  for  $W_{high}^k$ 
IF Execution Start Then
    Enforce (Action $_{app}^k$  (app properties)) to  $W_{high}^k$ 
    WHILE Execution.
        IF  $N_p$  equals  $W_{low}^k$  THEN
             $P_{mid}^k =$  Enforce (Action $_{mid}^k$  (res properties)) To
            PQRM
            Result = Enforce (Action $_{mid}^k$  (wf properties))
            To  $W_{low}^k$ 
            Return Result $_{mid}^k$ 
        ELSE
            Run in  $N_p$ 
        END IF
    END WHILE
END IF
END
    
```

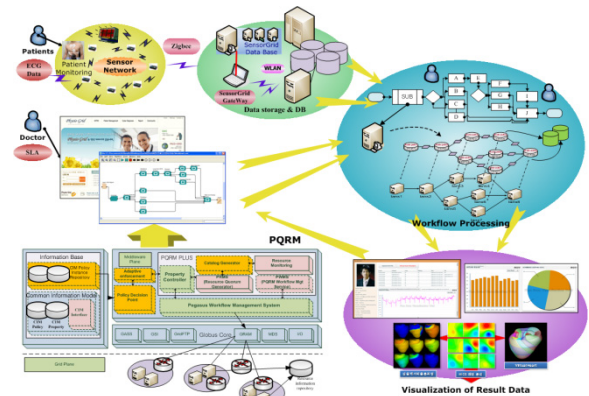
[그림4] 정책 결정 알고리즘

정책은 앞서 설명하였듯이 다음과 같이 정의 될 수 있다. 여기서 K는 사용자의 타입을 나타내며 app는 응용의 타입을 나타낸다. 이렇게 정의된 SLA의 각각의 정책은 응용, 워크플로우, 자원으로 나누어질 수 있는데 이를 시스템이 평가하기 위한 방법을 아래와 같다.

4. 시스템 평가

4.1. 실험 대상 : 심혈관계 진단 시스템

아래는 심혈관계 질병 진단을 하기 위한 프로세스를 나타낸다. 시스템은 원격의 맥내에 있는 환자로부터 ECG정보를 전송 받는다. 전송 받은 ECG는 일차적인 분석을 통해 질병에 대한 가능성이 존재하는 지를 확인하고 그러한 가능성이 존재하는 경우 가상심장을 실행하여 문제가 발생하는 심장을 확인할 수 있게 하여 준다. 그림 5에서는 심혈관계 질환 진단 시나리오를 실현한 시스템 예시를 보여준다.



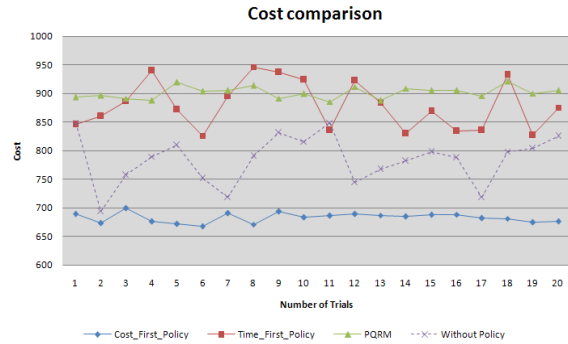
[그림5] 심혈관계 진단 시스템

4.2. 실험 환경

실험은 실험실에 있는 클러스터 자원 20대를 이용하여 수행되었다. 각각의 자원은 펜티엄4(1.4GHZ)이며 1Gbit의 스위치를 서로 연결되어 있다. 아래의 그림 6는 진단의 결과를 보여주기 위한 타일드 디스플레이라는 시스템을 나타내며 각각의 화면에 웹포털을 통해 진단 결과인 가상 심장, 그리고 ECG 처리 결과를 확인할 있다. 아래의 4가지 모니터는 실제 진단 서비스가 수행되는 프로세스를 나타내며 차례대로 실시간 ECG Monitoring, 환자선택, 환자 ECG 분석, 가상심장 정보를 나타낸다.



[그림6] 실험실에 구축된 심혈관계 진단 시스템

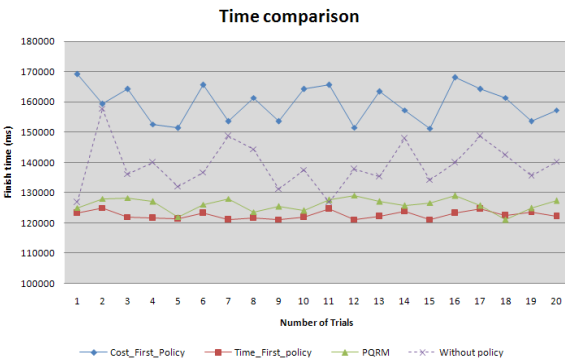


[그림8] 수행 비용 정책 기반 실험

4.3. 실험 결과

4.3.1. 완료시간 정책 기반 워크플로우 실행

워크플로우 완료 시간 정책은 워크플로우 수행시 그 수행시간을 최소화하려 것이다 본 실험에서는 위에서 정의된 심혈관계 질환을 위한 워크플로우를 가지고 2가지 정책(완료시간 우선, 수행비용 우선)과PQRM만을 이용하여 자원을 할당하는 경우, 그리고 랜덤으로 자원을 할당하는 경우 비교하였다. 아래의 그래프에서 보듯이 빨간색 점선은 완료시간을 우선시 하는 경우로써 4가지 경우 중에서 가장 수행시간이 적게 걸리는 것을 확인할 수 있다.



[그림7] 완료 시간 정책 기반 실험

4.3.1. 수행 비용 정책 기반 워크플로우 실행

워크플로우 수행 비용 정책은 워크플로우 수행시 그 수행 비용을 최소화하려 것이다 본 실험에서는 위에서 정의된 심혈관계 질환을 위한 워크플로우를 가지고 2가지 정책(완료시간 우선, 수행비용 우선)과PQRM만을 이용하여 자원을 할당하는 경우, 그리고 랜덤으로 자원을 할당하는 경우 비교하였다. 아래의 그래프에서 보듯이 빨간색 점선은 수행 비용을 우선시 하는 경우로써 4가지 경우 중에서 가장 수행 비용이 적게 걸리는 것을 확인할 수 있다.

5. 결론

본 논문에서는 심혈관계 진단 시스템에 정책그리드 기반의 워크플로우를 적용하기 위해 그리드 워크플로우의 정책과 기 개발된 그리드 자원관리 정책과의 통합방법을 제시하였으며 이를 위한 시스템 구조 및 알고리즘을 제안하였다. 시스템 구조적으로 정책 조정자를 통해 기 개발되었던 PQRM시스템과 새로 개발된 그리드 워크플로우 사이의 유기적인 통합을 이루었다. 본 논문은 이러한 정책기반의 워크플로우를 심혈관계 진단 시스템에 적용 결과 진단의 시간과 비용이 줄어들음을 실제 실험을 통해 확인하였다.

감사의 글

본 연구는 지식경제부가 지원하는 IT R&D 프로그램의 지원을 받아 수행되었습니다. [2008-C1090-0801-0014 and 2008-F-029-0

참고문헌

- [1] Eun Bo Shim et al., "Development of a 3D virtual heart including the mechanism of human cardiac cell physiology-electric conduction of cardiac tissue-heart mechanics", The Korean Society of Mechanical Engineers 60th anniversary fall conference 2005, The Korean Society of Mechanical Engineers, pp58-58, 2005.
- [2] K Krauter, R Buyya, M Maheswaran, "Taxonomy of Resource Management in Grid", Software Practice and Experience, 2002
- [3] T. Oinn, M. Addis, J. Ferris, D. Marvin, M. Senger, M. Greenwood, T. Carver and K. Glover, M.R. Pocock, A. Wipat, and P. Li, "Taverna: a tool for the composition and enactment of bioinformatics workflows". Bioinformatics, 20(17):3045-3054, Oxford University Press, London, UK, 2004.
- [4] Ian J. Taylor et al., "Workflows for e-Science", chap. 13. Oxford: Clarendon, 1892, pp.190-191.
- [5] K Yang, A Galis, C Todd, "Policy-based active Grid management architecture", ICON 2002. 10th IEEE International
- [6] Dong Su Nam et al., "QoS-constraint Quorum Management scheme for a Grid-based Electrocardiogram", LNCS 3043 pp.352-359, 2004.